

Ephemeral Hadoop Clusters in the Cloud



[1]

Greg Fodor, Etsy

gfodor@etsy.com

about me

gfodor@etsy.com

@gfodor

Data Wrangler

about etsy

the world's handmade marketplace

Buy Sell Community Blog Help

Etsy Register Sign In Handmade ▾

Categories

- Accessories
- Art
- Bags and Purses
- Bath and Beauty
- Books and Zines
- Candles
- Ceramics and Pottery
- Children
- Clothing
- Crochet
- Dolls and Miniatures
- Everything Else
- Furniture
- Geekery
- Glass
- Holidays
- Housewares
- Jewelry
- Knitting
- Music
- Needlecraft
- Paper Goods
- Patterns
- Pets
- Plants and Edibles
- Quilts
- Supplies
- Toys
- Vintage
- Weddings
- Woodworking

Handpicked Items See more

Picked by DirdyBirdy

 Water Nymph - Sugar Starfi... dreamsbbyteres... \$30.50 USD	 Hand Crocheted Mini Bows i... cornflowerbl... \$4.00 USD	 Vintage Collection of Bask... DeLaBelle \$22.00 USD	 Orostachys Succulent In A ... airyobsessio... \$4.50 USD
 Renewal vintage material s... kellyemerald... \$6.00 USD	 14k Solid Rose Gold Lightn... TheAngryWeat... \$159.00 USD	 Vintage Sea Coral Print - ... vintagebythe... \$14.00 USD	 Oh Crab! - Hand Embroidered... deepindigo \$7.32 USD
 Patchouli Soap morganstreet \$6.00 USD	 Blue Chevron Tea Towel erindollar \$18.00 USD	 Sea Shells by the Sea Shor... Birdnestsand... \$4.50 USD	 Twisted cord green wrap br... Myknottin... \$16.00 USD

total members: 9,000,000

total active shops: 800,000

items listed: 9.5M

page views per month: >1B

2010 sales: \$314.3M

lots of data

about this talk

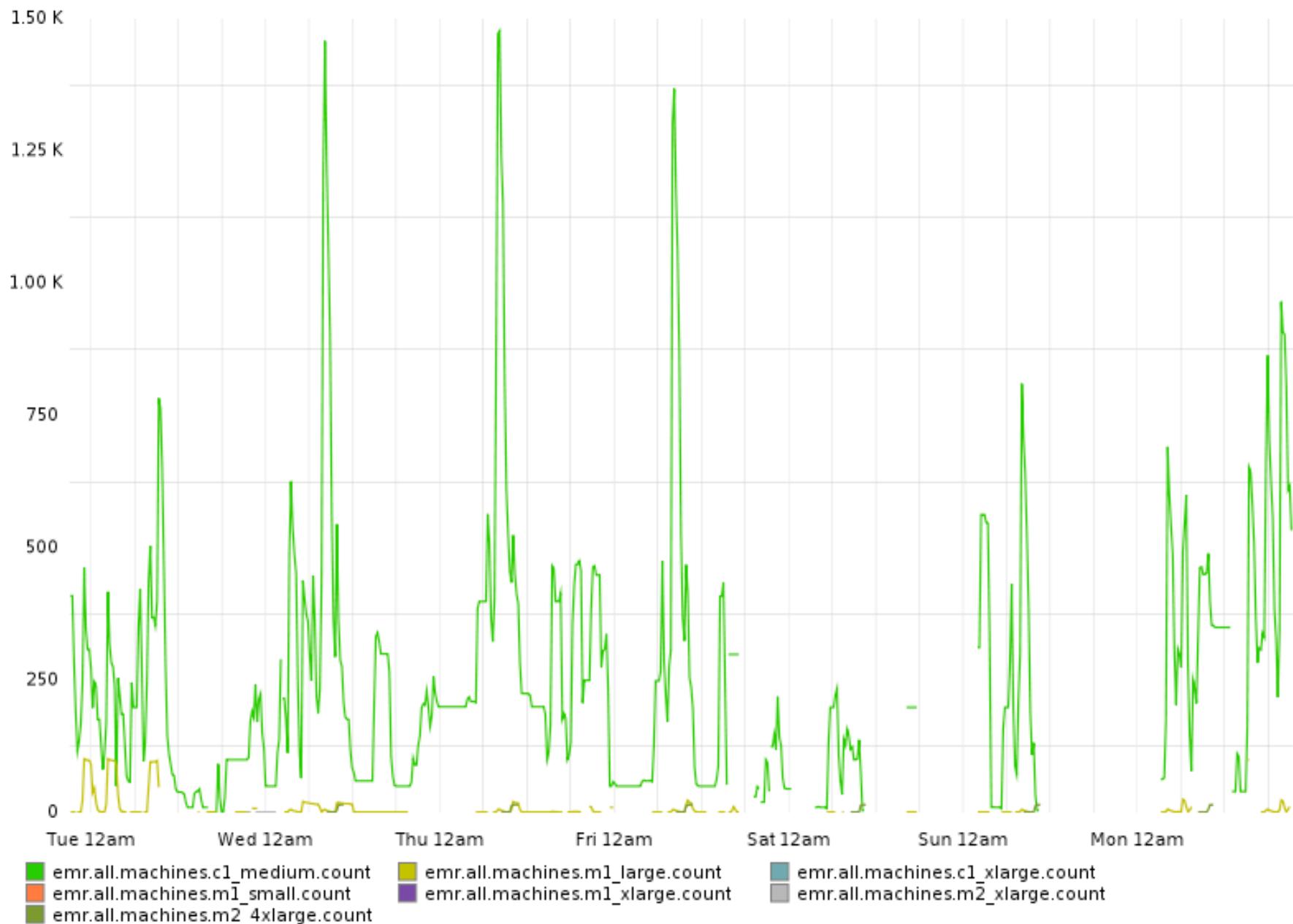
ephemeral?

[5]

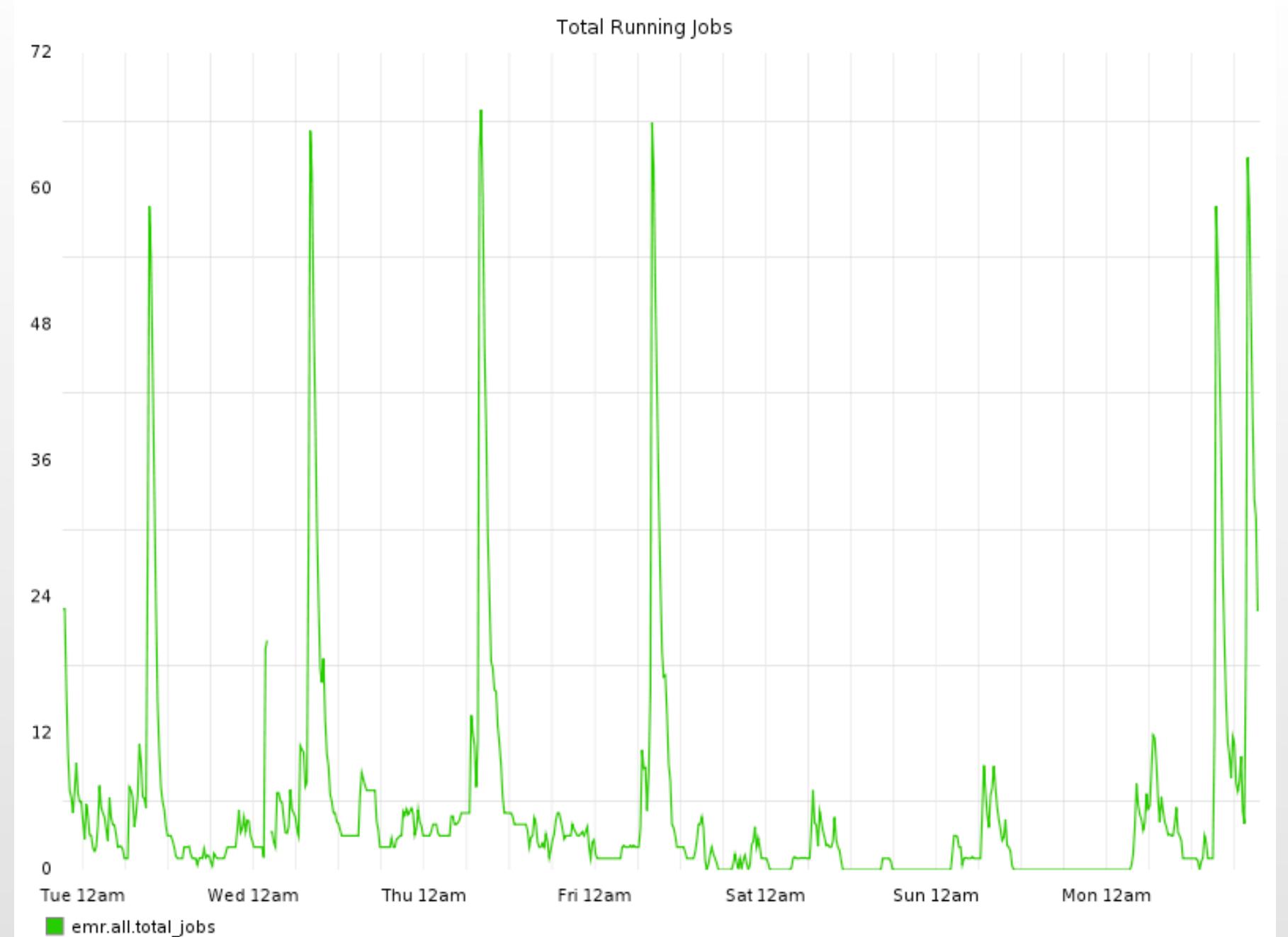


“elastic” to the extreme

Machine breakdown







how did we get here?

wanted to dip our toes

stop hitting the database

stop grepping log files

2 data sources -> S3

database snapshots

input:
nightly diffs

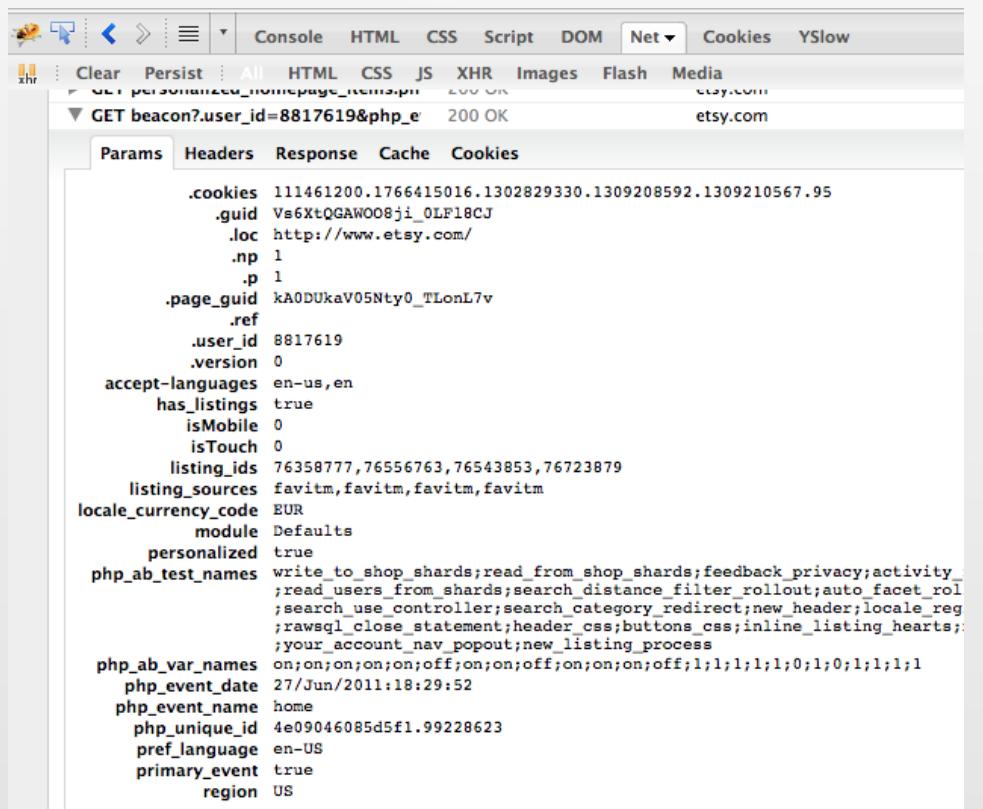
(SELECT * FROM <table> WHERE update_date > 1 day ago)

output:
full tables as sequence files

visit logs

input:
akamai access logs
(event beacons)

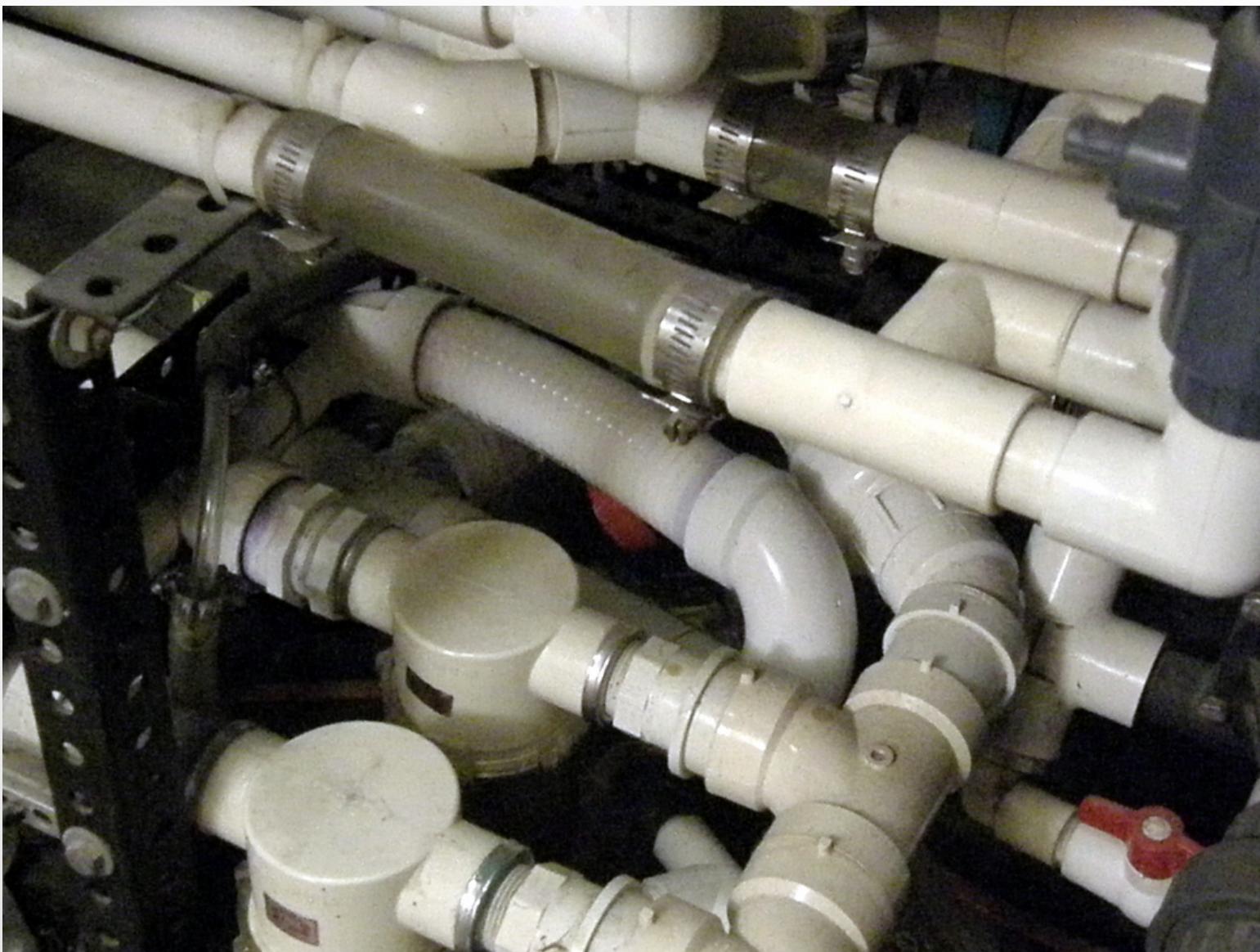
output:
[visit_id, [event]]



The screenshot shows a browser's developer tools Network tab with a single request listed:

- Method: GET
- URL: beacon?user_id=8817619&php_e
- Status: 200 OK
- Headers (partial): XHR, Images, Flash, Media, etsy.com
- Params (selected tab):
 - .cookies: 111461200.1766415016.1302829330.1309208592.1309210567.95
 - .guid: Vs6XtQGAW008ji_0LF18CJ
 - .loc: http://www.etsy.com/
 - .np: 1
 - .p: 1
 - .page_guid: kA0DUkaV05Nty0_TLonL7v
 - .ref:
 - .user_id: 8817619
 - .version: 0
 - accept-languages: en-us,en
 - has_listings: true
 - isMobile: 0
 - isTouch: 0
 - listing_ids: 76358777,76556763,76543853,76723879
 - listing_sources: favitm,favitm,favitm,favitm
 - locale_currency_code: EUR
 - module: Defaults
 - personalized: true
 - php_ab_test_names: write_to_shop_shards;read_from_shop_shards;feedback_privacy;activity_;read_users_from_shards;search_distance_filter_rollout;auto_facet rol;search_use_controller;search_category_redirect;new_header;locale_reg;rawsql_close_statement;header_css;buttons_css;inline_listing_hearts;your_account_nav_popup;new_listing_process
 - php_ab_var_names: on;on;on;on;on;off;on;on;off;on;on;on;on;off;1;1;1;1;1;0;1;0;1;1;1;1
 - php_event_date: 27/Jun/2011:18:29:52
 - php_event_name: home
 - php_unique_id: 4e09046085d5f1.99228623
 - pref_language: en-US
 - primary_event: true
 - region: US

processing the data



[2]

data flow

joins, group bys, etc.

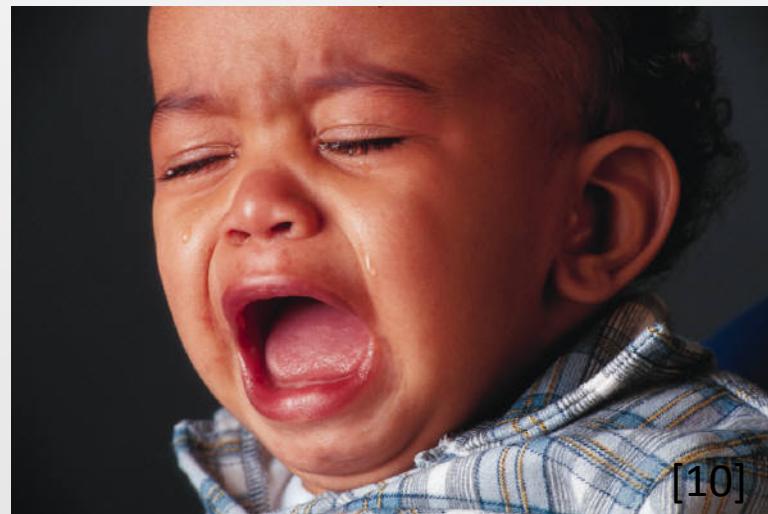
cascading

Chris Wensel

<http://www.cascading.org/>

great implementation

Java syntax



cascading.jruby

Grégoire Marabout (Qualtera), Matt Walker (Etsy), Stefan Karpinski (Etsy), Steve Mardenfeld (Etsy)

github: <http://bit.ly/o3DNtC>
blog: <http://etsy.me/cFytuL>

```
cascade 'tag_counter', :input_prefix, :output_prefix do
  flow 'tag counter' do
    tap_db_snapshot "listings"

    assembly "listings" do
      split_rows "tags", "tag",
                  :pattern => /\./,
                  :output => ["listing_id", "tag"]

      branch "tag_counts" do
        group_by "tag" do
          count "tag_count"
        end
      end
    end

    sink 'tag_counts', :scheme => text_line_scheme
  end
end
```

“push” job binaries to S3

run on Elastic Map/Reduce

starts cluster, runs, shuts down

access results on S3

next project:
shop recommendations

3 steps:

- ✓ data preparation - Cascading
- ✗ analysis/training
- ✗ prediction



sparse implementation of SVD

3 steps:

- ✓ data preparation - Cascading
- ✗ analysis/training - MATLAB
- ✗ prediction - MATLAB



X

“MATLAB, in *my*
Hadoop cluster?”

It's more likely than you think.

FREE PC CHECK!



CONTENTwatch™

hadoop streaming

arbitrary scripts for map & reduce

Swiss army knife

Full dataset analysis

Matlab, Ruby scripts

'Artifact' outputs

Tokyo Cabinet, Lucene, SQLite

Side-effects

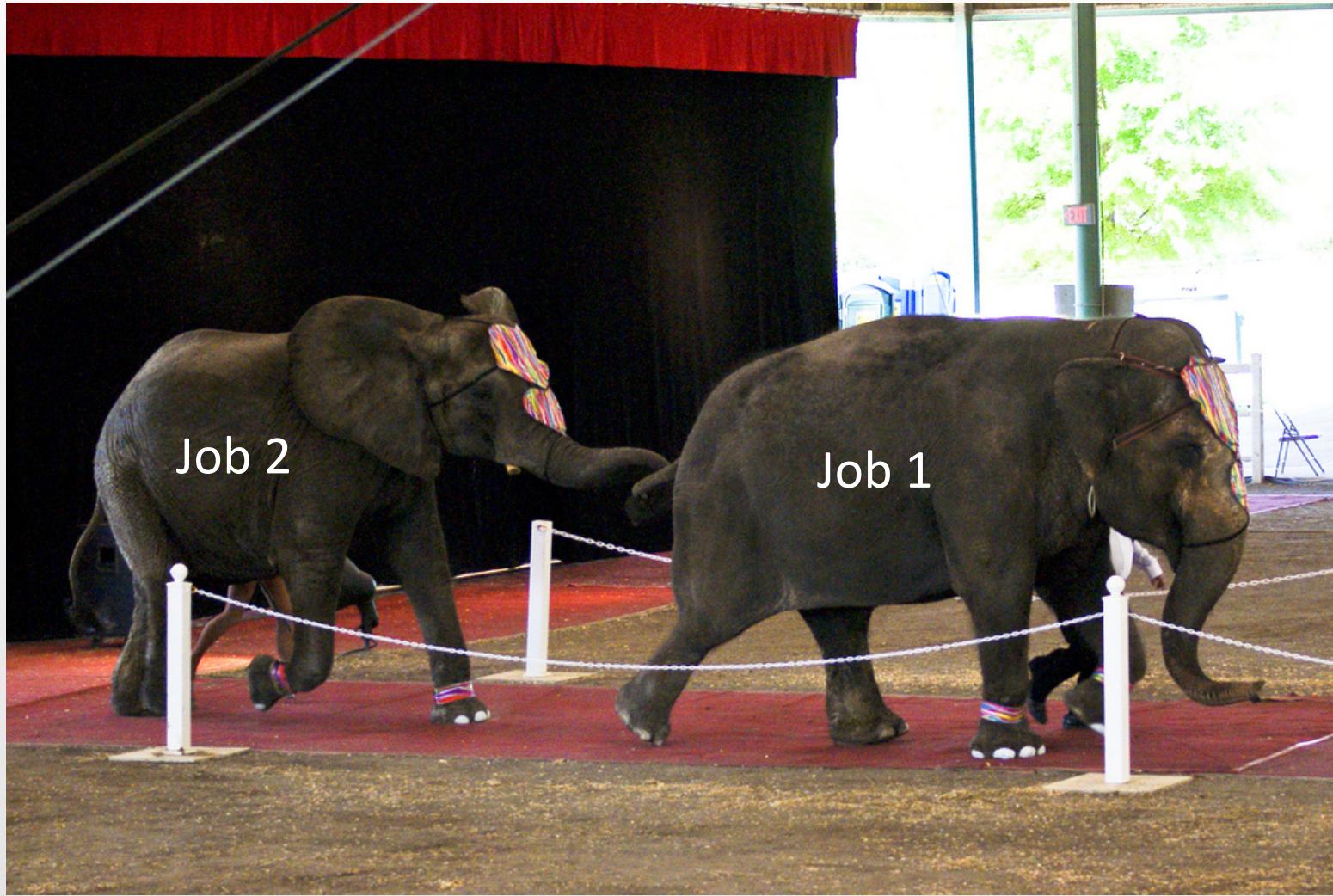
MySQL, CloudFront



[3]

3 steps:

- ✓ data preparation - Cascading
- ✓ analysis/training - MATLAB
- ✓ prediction - MATLAB



[4]

Barnum

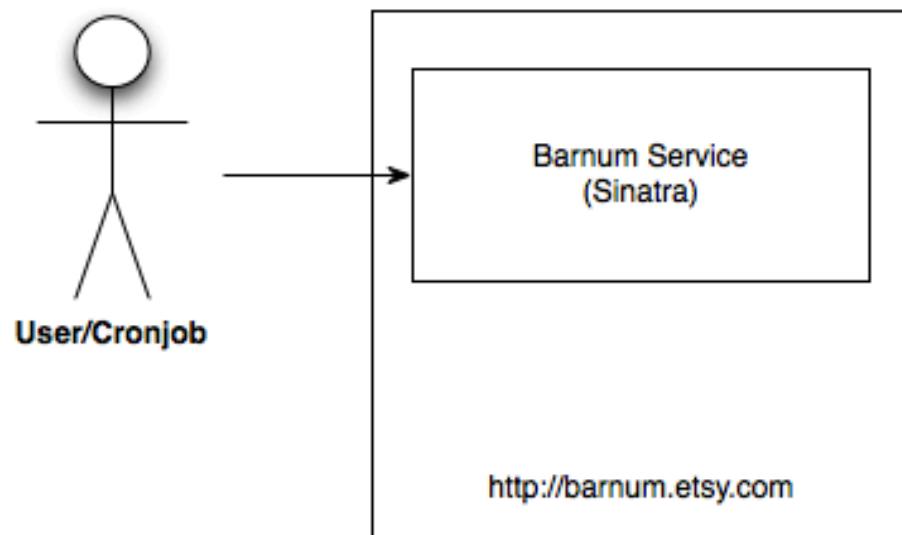
Sinatra web service on EC2

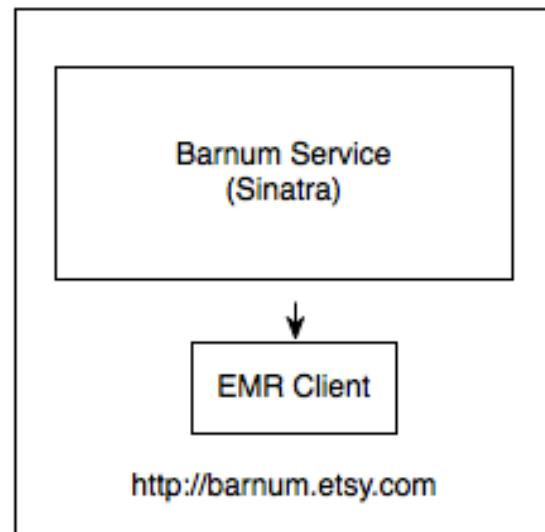
barnum starts job and passes
callback URL

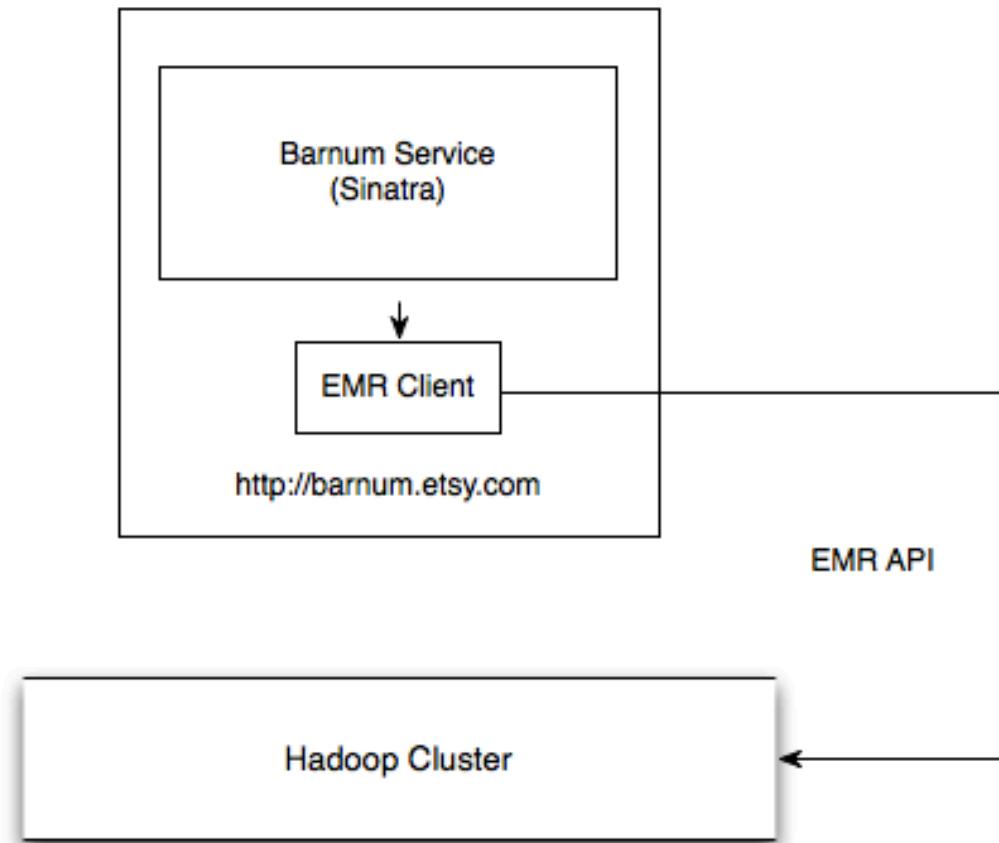
when job finishes, hadoop hits
callback URL to barnum to proceed

Barnum Service
(Sinatra)

<http://barnum.etsy.com>



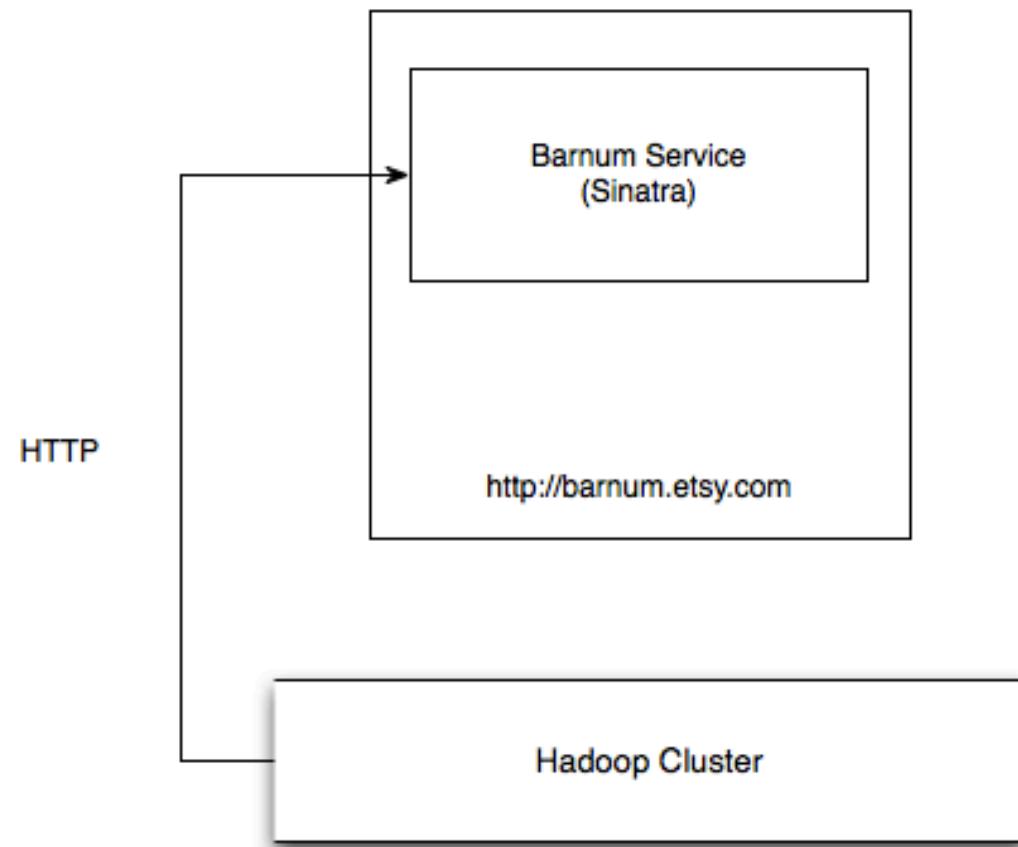


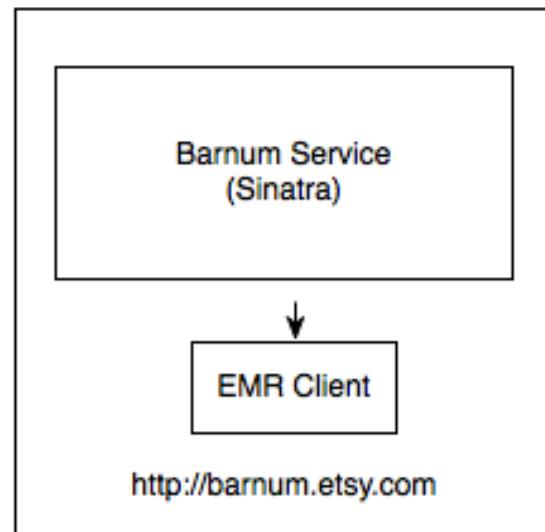


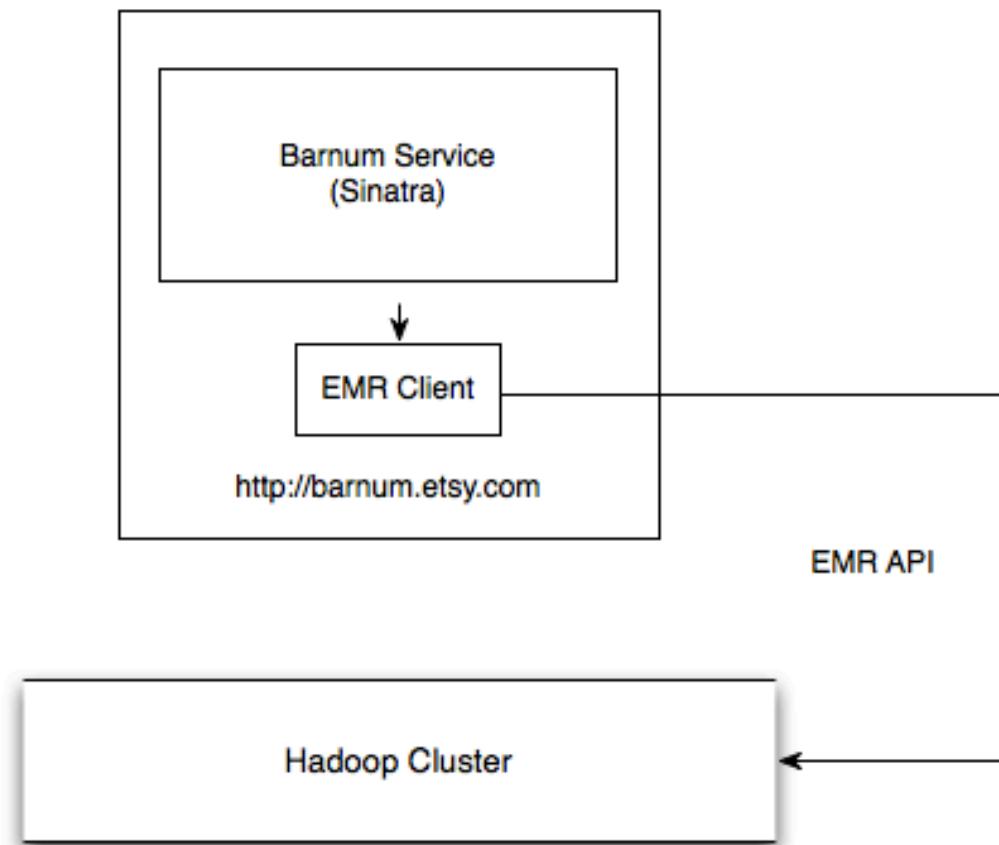
Barnum Service
(Sinatra)

<http://barnum.etsy.com>

Hadoop Cluster







Barnum constructs

Task 1

Job A

Cascading
45 c1.mediums

Task 1

Job A

Cascading
45 c1.mediums

Task 2

Job B

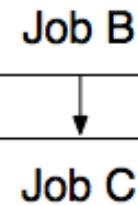
Job C

Chaining

Task 1

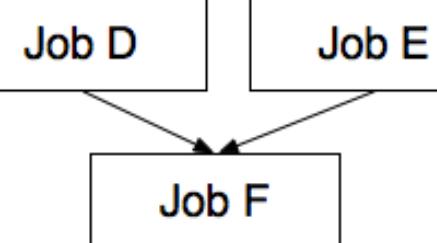
Job A
Cascading
45 c1.mediums

Task 2



Chaining

Task 3

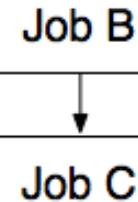


Batching

Task 1

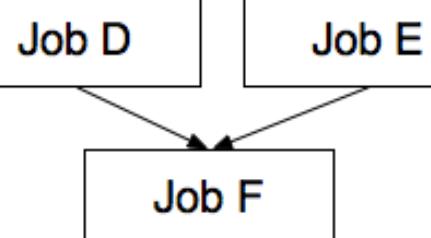
Job A
Cascading
45 c1.mediums

Task 2



Chaining

Task 3



Batching

Task 1

Task 2

Task 3

Branching

3 steps:

- ✓ data preparation - Cascading
- ✓ analysis/training - MATLAB
- ✓ prediction - MATLAB

suggested_shops.yaml:

```
schedule: 15 07 * * *

transitions:
  START:
    SUCCEEDED: joinFavorites
  joinFavorites:
    SUCCEEDED: analysis
  analysis:
    SUCCEEDED: prediction
  prediction:
    SUCCEEDED: DONE

params:
  ALL:
    data_prefix:      s3://data.etsy.com
    taste_prefix:     s3://taste.etsy.com
  joinFavorites:
    instance_type:   c1.medium
    instance_count:  20
    min_listings:    3
  analysis:
    instance_type:   m2.4xlarge
    instance_count:  1
  prediction:
    instance_type:   m2.4xlarge
    instance_count:  10
    recommendations: 200
```

User table snapshot (S3)

Favorites table snapshot (S3)

suggested_shops.yaml:

```
schedule: 15 07 * * *

transitions:
  START:
    SUCCEEDED: joinFavorites
  joinFavorites:
    SUCCEEDED: analysis
  analysis:
    SUCCEEDED: prediction
  prediction:
    SUCCEEDED: DONE

params:
  ALL:
    data_prefix:      s3://data.etsy.com
    taste_prefix:     s3://taste.etsy.com
  joinFavorites:
    instance_type:   c1.medium
    instance_count:  20
    min_listings:    3
  analysis:
    instance_type:   m2.4xlarge
    instance_count:  1
  prediction:
    instance_type:   m2.4xlarge
    instance_count:  10
    recommendations: 200
```

User table snapshot (S3)

Favorites table snapshot (S3)

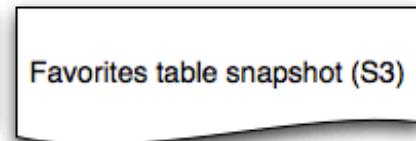
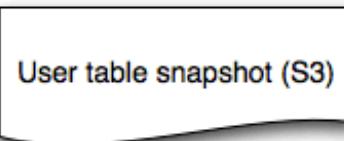
joinFavorites(min_listings:3)
Cascading
20 c1.mediums

suggested_shops.yaml:

```
schedule: 15 07 * * *

transitions:
  START:
    SUCCEEDED: joinFavorites
  joinFavorites:
    SUCCEEDED: analysis
  analysis:
    SUCCEEDED: prediction
  prediction:
    SUCCEEDED: DONE

params:
  ALL:
    data_prefix: s3://data.etsy.com
    taste_prefix: s3://taste.etsy.com
  joinFavorites:
    instance_type: c1.medium
    instance_count: 20
    min_listings: 3
  analysis:
    instance_type: m2.4xlarge
    instance_count: 1
  prediction:
    instance_type: m2.4xlarge
    instance_count: 10
    recommendations: 200
```



joinFavorites(min_listings:3)
Cascading
20 c1.mediums

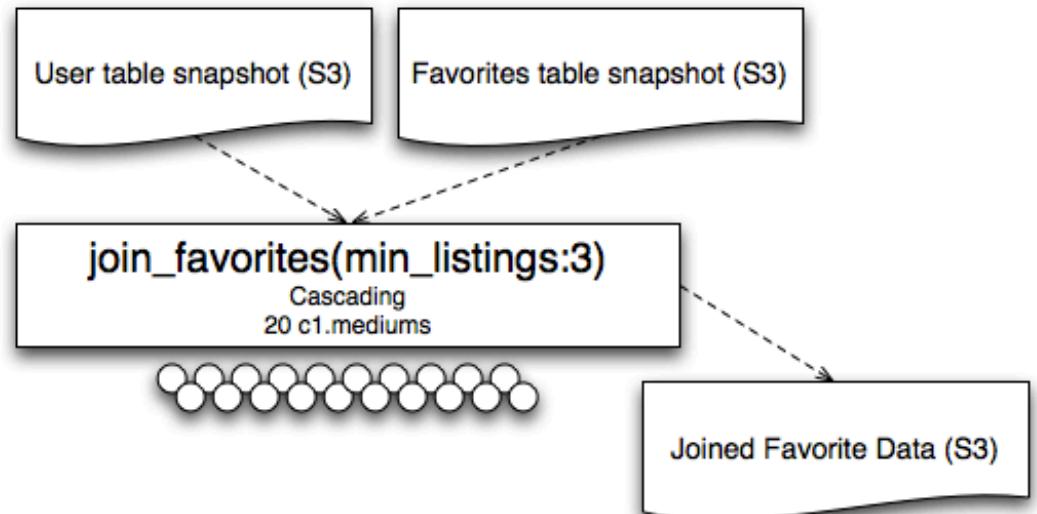


suggested_shops.yaml:

```
schedule: 15 07 * * *

transitions:
  START:
    SUCCEEDED: joinFavorites
  joinFavorites:
    SUCCEEDED: analysis
  analysis:
    SUCCEEDED: prediction
  prediction:
    SUCCEEDED: DONE

params:
  ALL:
    data_prefix: s3://data.etsy.com
    taste_prefix: s3://taste.etsy.com
  joinFavorites:
    instance_type: c1.medium
    instance_count: 20
    min_listings: 3
  analysis:
    instance_type: m2.4xlarge
    instance_count: 1
  prediction:
    instance_type: m2.4xlarge
    instance_count: 10
    recommendations: 200
```



suggested_shops.yaml:

```
schedule: 15 07 * * *

transitions:
  START:
    SUCCEEDED: joinFavorites
  joinFavorites:
    SUCCEEDED: analysis
  analysis:
    SUCCEEDED: prediction
  prediction:
    SUCCEEDED: DONE

params:
  ALL:
    data_prefix:      s3://data.etsy.com
    taste_prefix:     s3://taste.etsy.com
  joinFavorites:
    instance_type:   c1.medium
    instance_count:  20
    min_listings:    3
  analysis:
    instance_type:   m2.4xlarge
    instance_count:  1
  prediction:
    instance_type:   m2.4xlarge
    instance_count:  10
    recommendations: 200
```



suggested_shops.yaml:

```
schedule: 15 07 * * *

transitions:
  START:
    SUCCEEDED: joinFavorites
  joinFavorites
    SUCCEEDED: analysis
  analysis:
    SUCCEEDED: prediction
  prediction:
    SUCCEEDED: DONE

params:
  ALL:
    data_prefix: s3://data.etsy.com
    taste_prefix: s3://taste.etsy.com
  joinFavorites:
    instance_type: c1.medium
    instance_count: 20
    min_listings: 3
  analysis:
    instance_type: m2.4xlarge
    instance_count: 1
  prediction:
    instance_type: m2.4xlarge
    instance_count: 10
    recommendations: 200
```

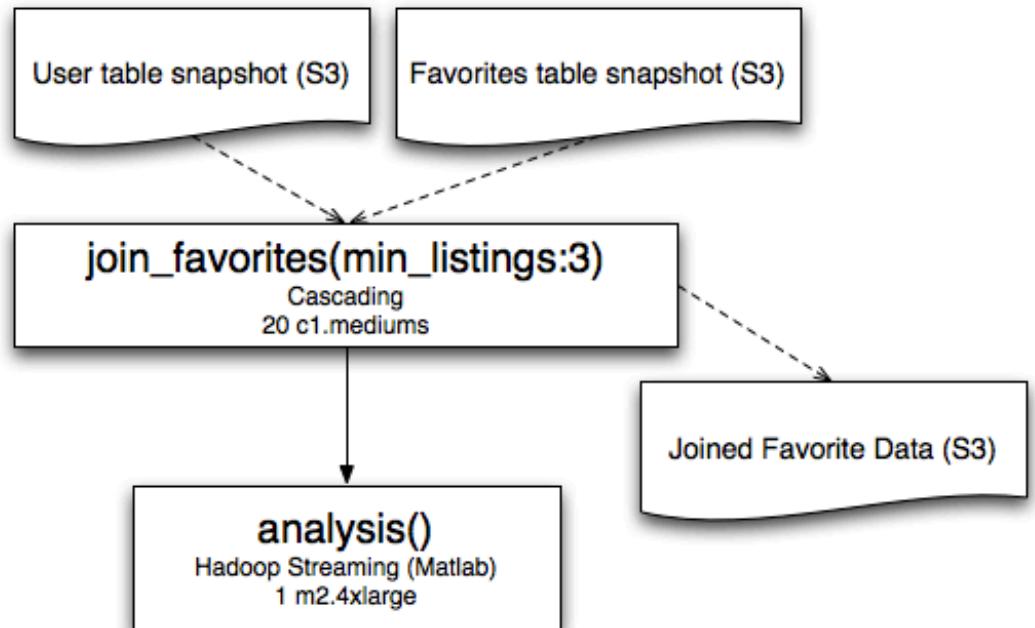


suggested_shops.yaml:

```
schedule: 15 07 * * *

transitions:
  START:
    SUCCEEDED: joinFavorites
  joinFavorites:
    SUCCEEDED: analysis
  analysis:
    SUCCEEDED: prediction
  prediction:
    SUCCEEDED: DONE

params:
  ALL:
    data_prefix: s3://data.etsy.com
    taste_prefix: s3://taste.etsy.com
  joinFavorites:
    instance_type: c1.medium
    instance_count: 20
    min_listings: 3
  analysis:
    instance_type: m2.4xlarge
    instance_count: 1
  prediction:
    instance_type: m2.4xlarge
    instance_count: 10
    recommendations: 200
```

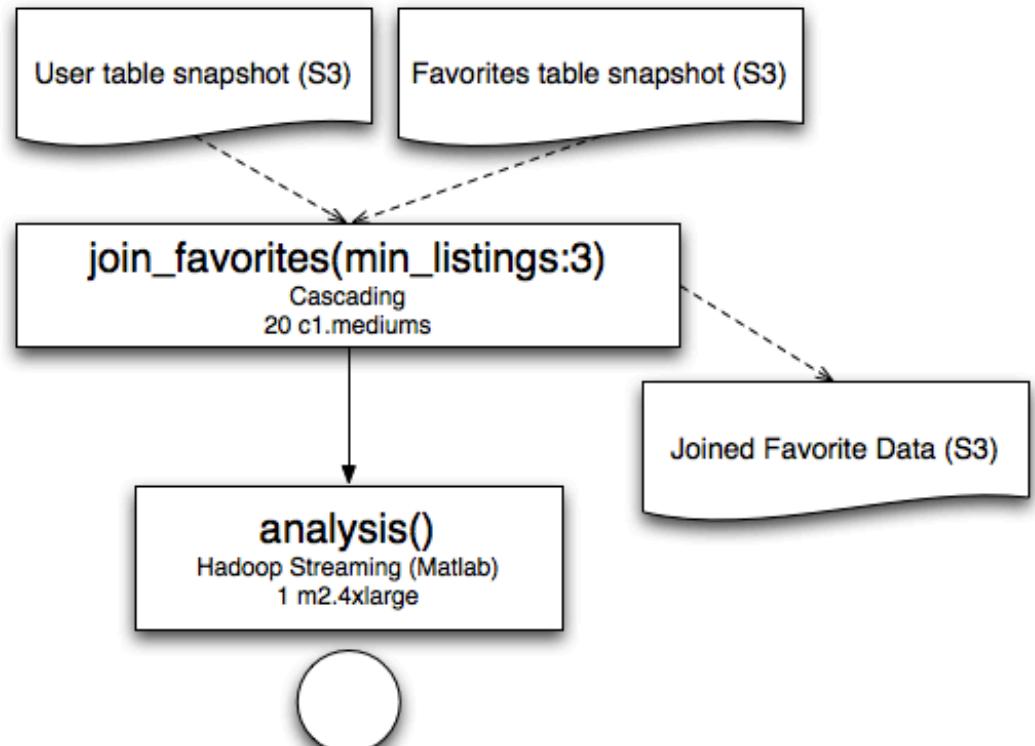


suggested_shops.yaml:

```
schedule: 15 07 * * *

transitions:
  START:
    SUCCEEDED: joinFavorites
  joinFavorites:
    SUCCEEDED: analysis
  analysis:
    SUCCEEDED: prediction
  prediction:
    SUCCEEDED: DONE

params:
  ALL:
    data_prefix: s3://data.etsy.com
    taste_prefix: s3://taste.etsy.com
  joinFavorites:
    instance_type: c1.medium
    instance_count: 20
    min_listings: 3
  analysis:
    instance_type: m2.4xlarge
    instance_count: 1
  prediction:
    instance_type: m2.4xlarge
    instance_count: 10
    recommendations: 200
```

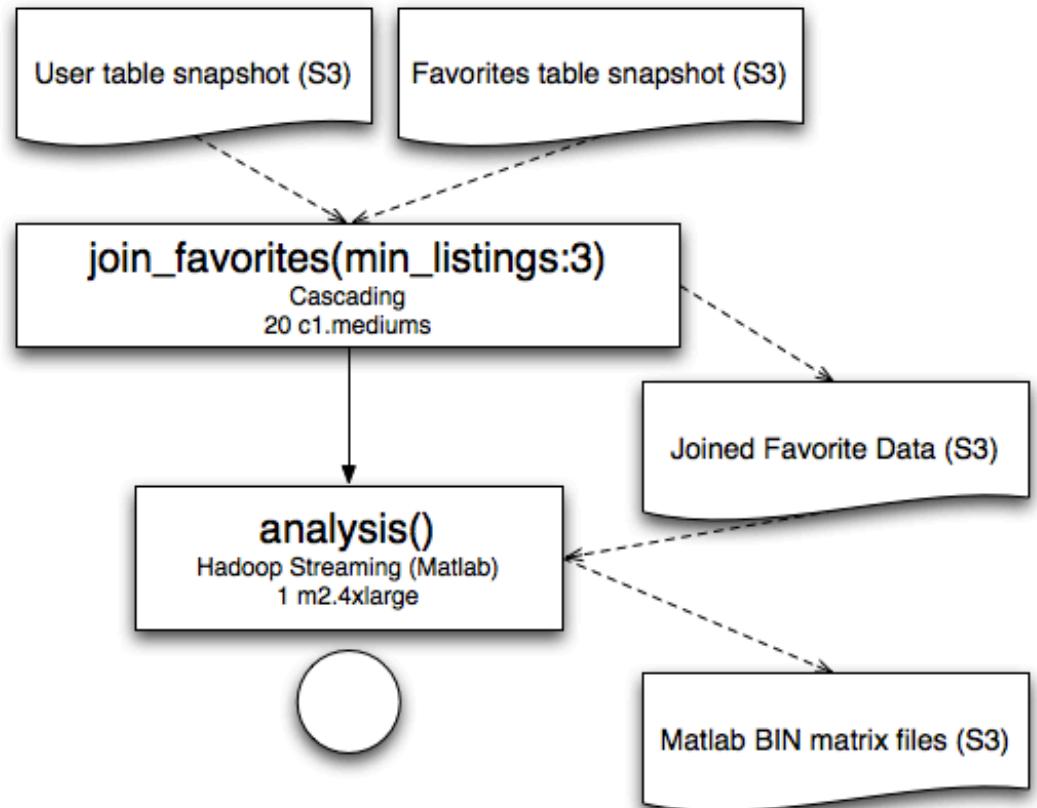


suggested_shops.yaml:

```
schedule: 15 07 * * *

transitions:
  START:
    SUCCEEDED: joinFavorites
  joinFavorites:
    SUCCEEDED: analysis
  analysis:
    SUCCEEDED: prediction
  prediction:
    SUCCEEDED: DONE

params:
  ALL:
    data_prefix: s3://data.etsy.com
    taste_prefix: s3://taste.etsy.com
  joinFavorites:
    instance_type: c1.medium
    instance_count: 20
    min_listings: 3
  analysis:
    instance_type: m2.4xlarge
    instance_count: 1
  prediction:
    instance_type: m2.4xlarge
    instance_count: 10
    recommendations: 200
```

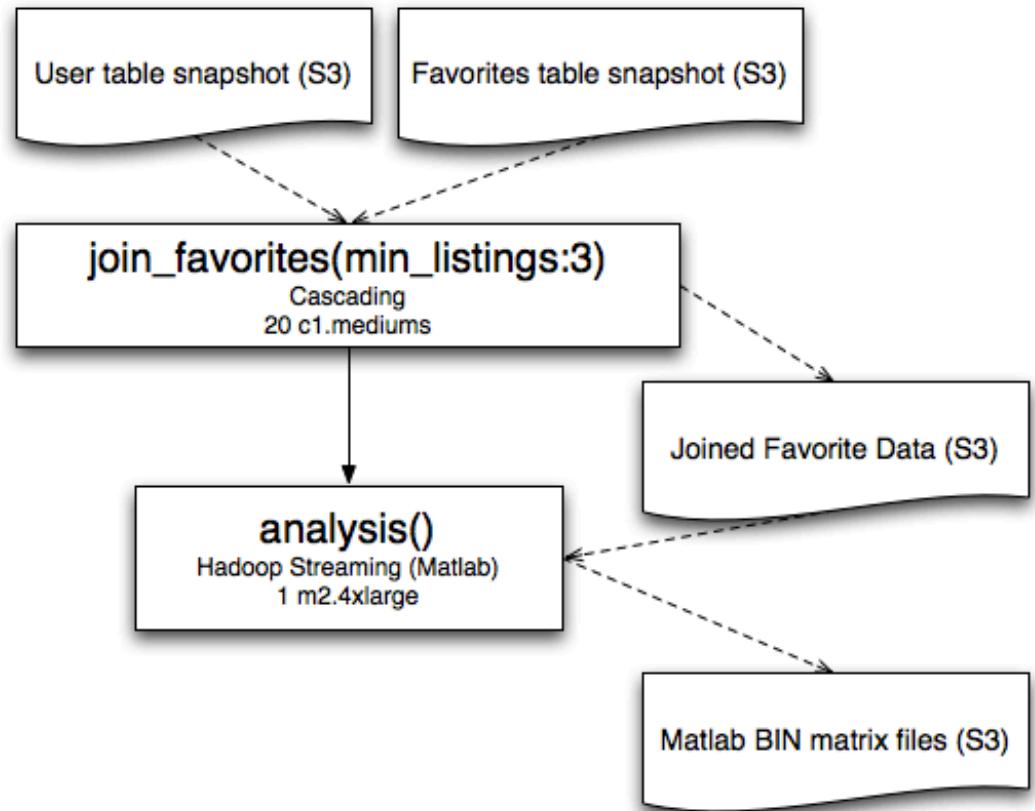


suggested_shops.yaml:

```
schedule: 15 07 * * *

transitions:
  START:
    SUCCEEDED: joinFavorites
  joinFavorites:
    SUCCEEDED: analysis
  analysis:
    SUCCEEDED: prediction
  prediction:
    SUCCEEDED: DONE

params:
  ALL:
    data_prefix: s3://data.etsy.com
    taste_prefix: s3://taste.etsy.com
  joinFavorites:
    instance_type: c1.medium
    instance_count: 20
    min_listings: 3
  analysis:
    instance_type: m2.4xlarge
    instance_count: 1
  prediction:
    instance_type: m2.4xlarge
    instance_count: 10
    recommendations: 200
```

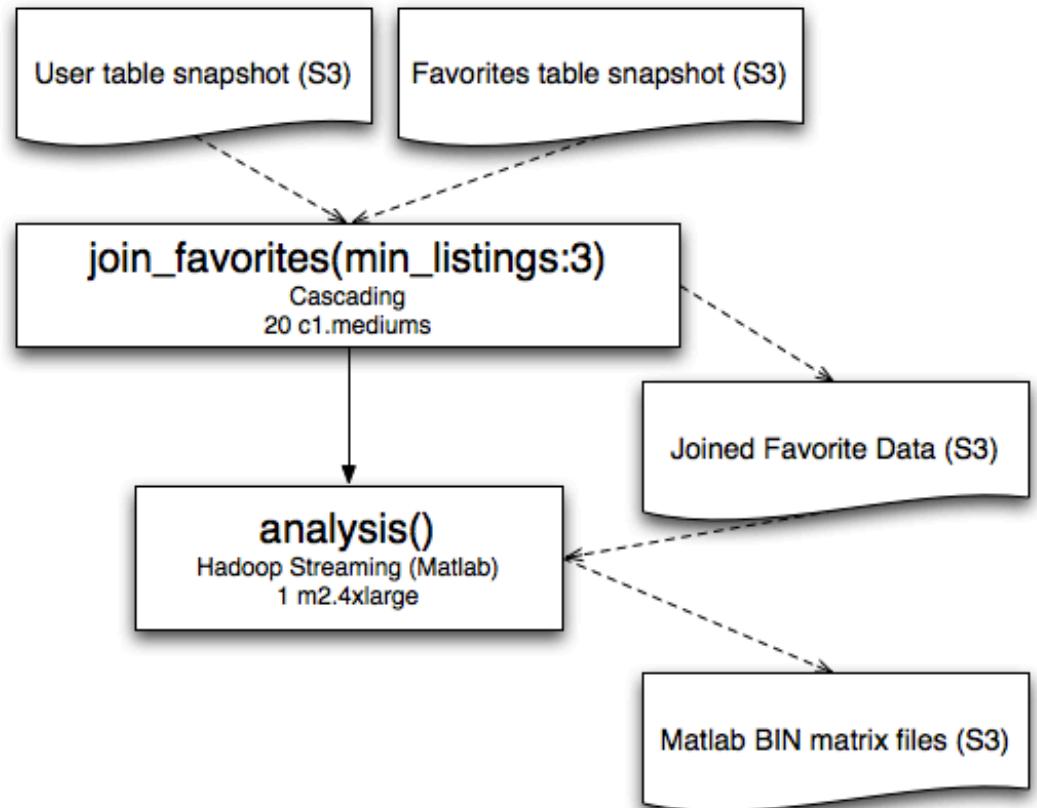


suggested_shops.yaml:

```
schedule: 15 07 * * *

transitions:
  START:
    SUCCEEDED: joinFavorites
  joinFavorites:
    SUCCEEDED: analysis
  analysis:
    SUCCEEDED: prediction
  prediction:
    SUCCEEDED: DONE

params:
  ALL:
    data_prefix: s3://data.etsy.com
    taste_prefix: s3://taste.etsy.com
  joinFavorites:
    instance_type: c1.medium
    instance_count: 20
    min_listings: 3
  analysis:
    instance_type: m2.4xlarge
    instance_count: 1
  prediction:
    instance_type: m2.4xlarge
    instance_count: 10
    recommendations: 200
```

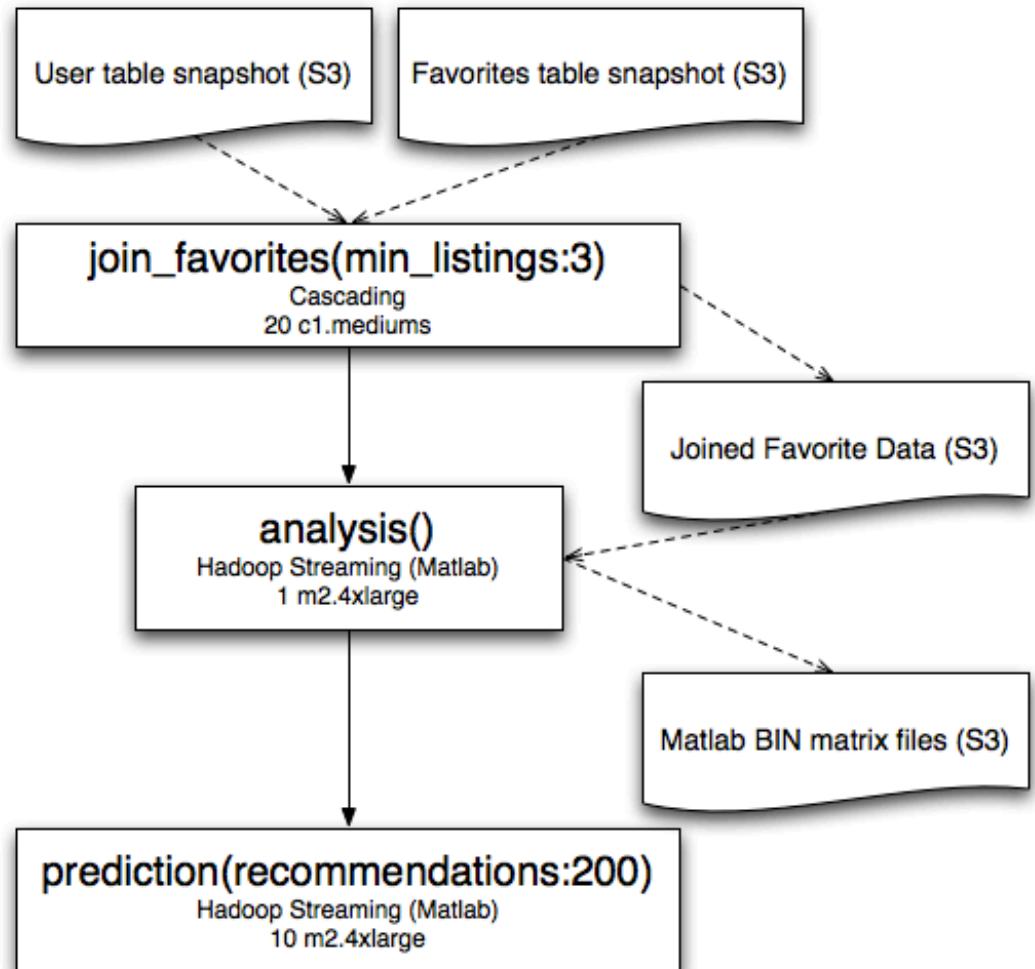


suggested_shops.yaml:

```
schedule: 15 07 * * *

transitions:
  START:
    SUCCEEDED: joinFavorites
  joinFavorites:
    SUCCEEDED: analysis
  analysis:
    SUCCEEDED: prediction
  prediction:
    SUCCEEDED: DONE

params:
  ALL:
    data_prefix: s3://data.etsy.com
    taste_prefix: s3://taste.etsy.com
  joinFavorites:
    instance_type: c1.medium
    instance_count: 20
    min_listings: 3
  analysis:
    instance_type: m2.4xlarge
    instance_count: 1
  prediction:
    instance_type: m2.4xlarge
    instance_count: 10
    recommendations: 200
```

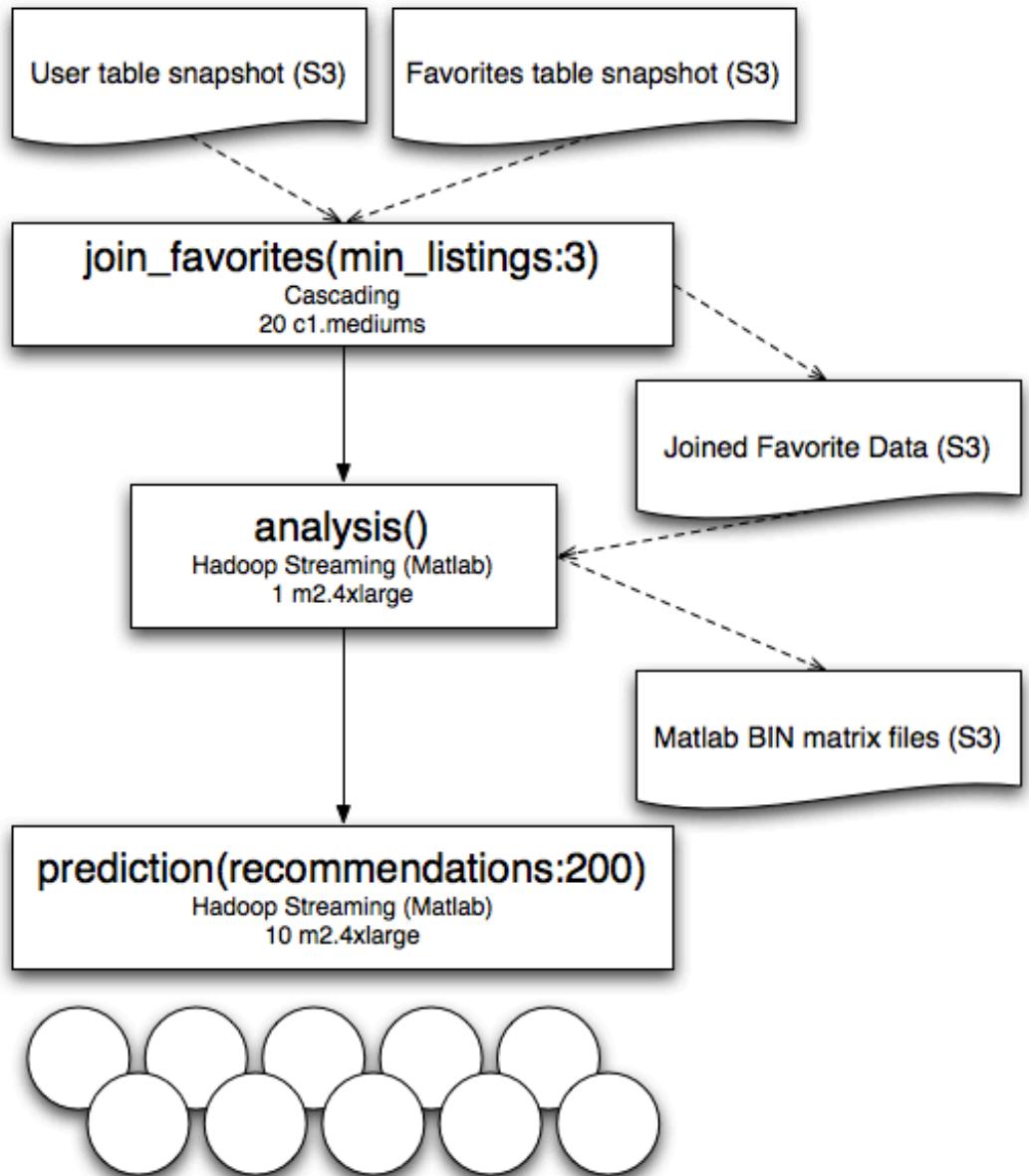


suggested_shops.yaml:

```
schedule: 15 07 * * *

transitions:
  START:
    SUCCEEDED: joinFavorites
  joinFavorites:
    SUCCEEDED: analysis
  analysis:
    SUCCEEDED: prediction
  prediction:
    SUCCEEDED: DONE

params:
  ALL:
    data_prefix: s3://data.etsy.com
    taste_prefix: s3://taste.etsy.com
  joinFavorites:
    instance_type: c1.medium
    instance_count: 20
    min_listings: 3
  analysis:
    instance_type: m2.4xlarge
    instance_count: 1
  prediction:
    instance_type: m2.4xlarge
    instance_count: 10
    recommendations: 200
```

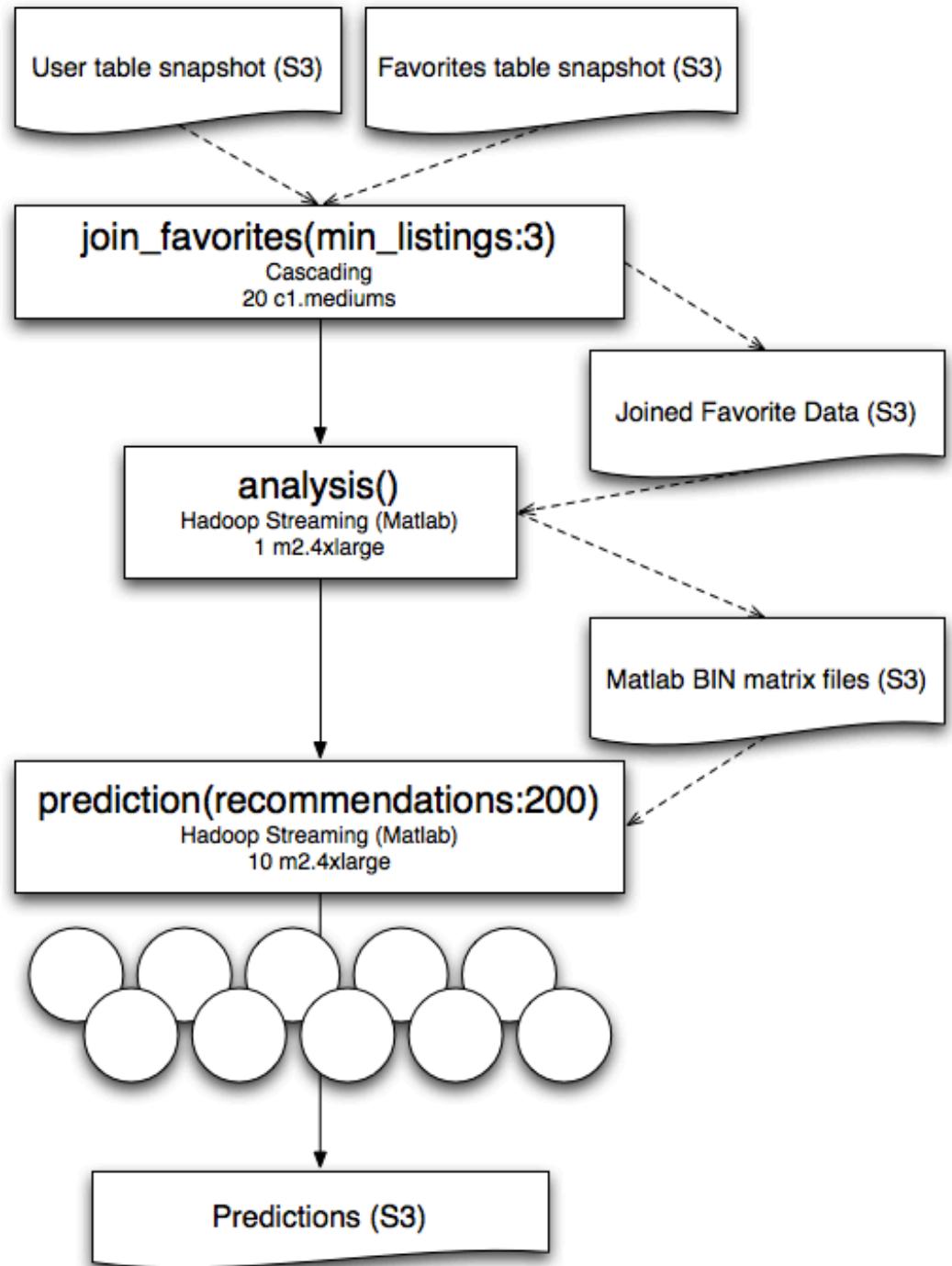


suggested_shops.yaml:

```
schedule: 15 07 * * *

transitions:
  START:
    SUCCEEDED: joinFavorites
  joinFavorites:
    SUCCEEDED: analysis
  analysis:
    SUCCEEDED: prediction
  prediction:
    SUCCEEDED: DONE

params:
  ALL:
    data_prefix: s3://data.etsy.com
    taste_prefix: s3://taste.etsy.com
  joinFavorites:
    instance_type: c1.medium
    instance_count: 20
    min_listings: 3
  analysis:
    instance_type: m2.4xlarge
    instance_count: 1
  prediction:
    instance_type: m2.4xlarge
    instance_count: 10
    recommendations: 200
```

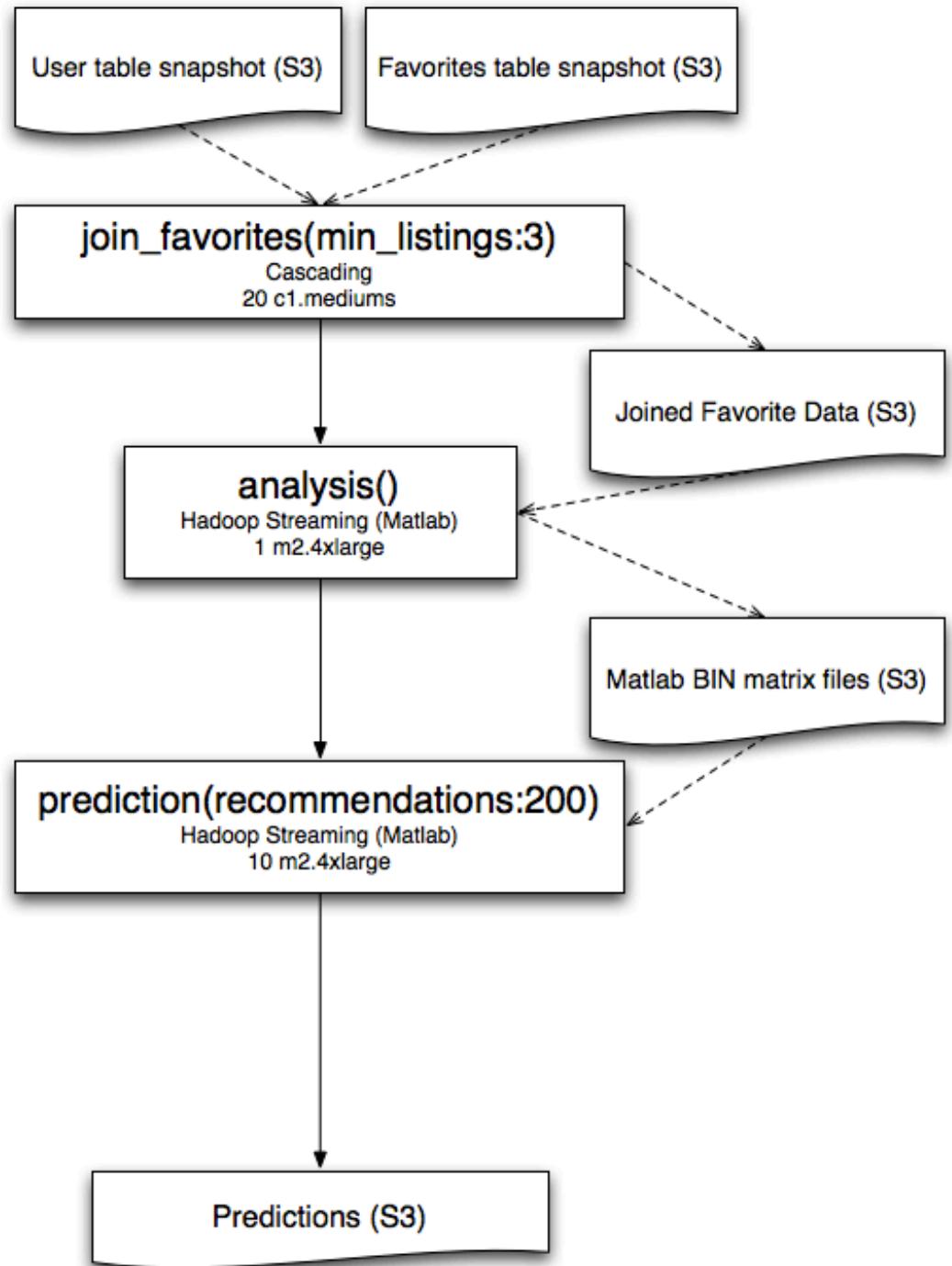


suggested_shops.yaml:

```
schedule: 15 07 * * *

transitions:
  START:
    SUCCEEDED: joinFavorites
  joinFavorites:
    SUCCEEDED: analysis
  analysis:
    SUCCEEDED: prediction
  prediction:
    SUCCEEDED: DONE

params:
  ALL:
    data_prefix: s3://data.etsy.com
    taste_prefix: s3://taste.etsy.com
  joinFavorites:
    instance_type: c1.medium
    instance_count: 20
    min_listings: 3
  analysis:
    instance_type: m2.4xlarge
    instance_count: 1
  prediction:
    instance_type: m2.4xlarge
    instance_count: 10
    recommendations: 200
```



getting data back to web stack?



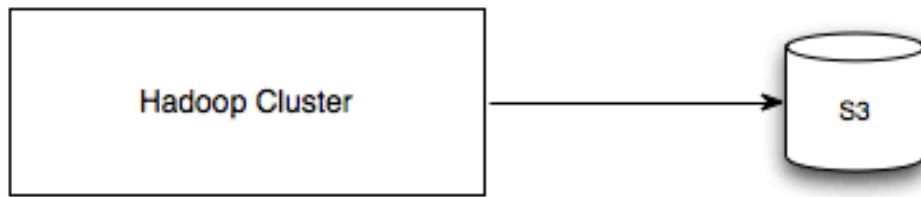
v1

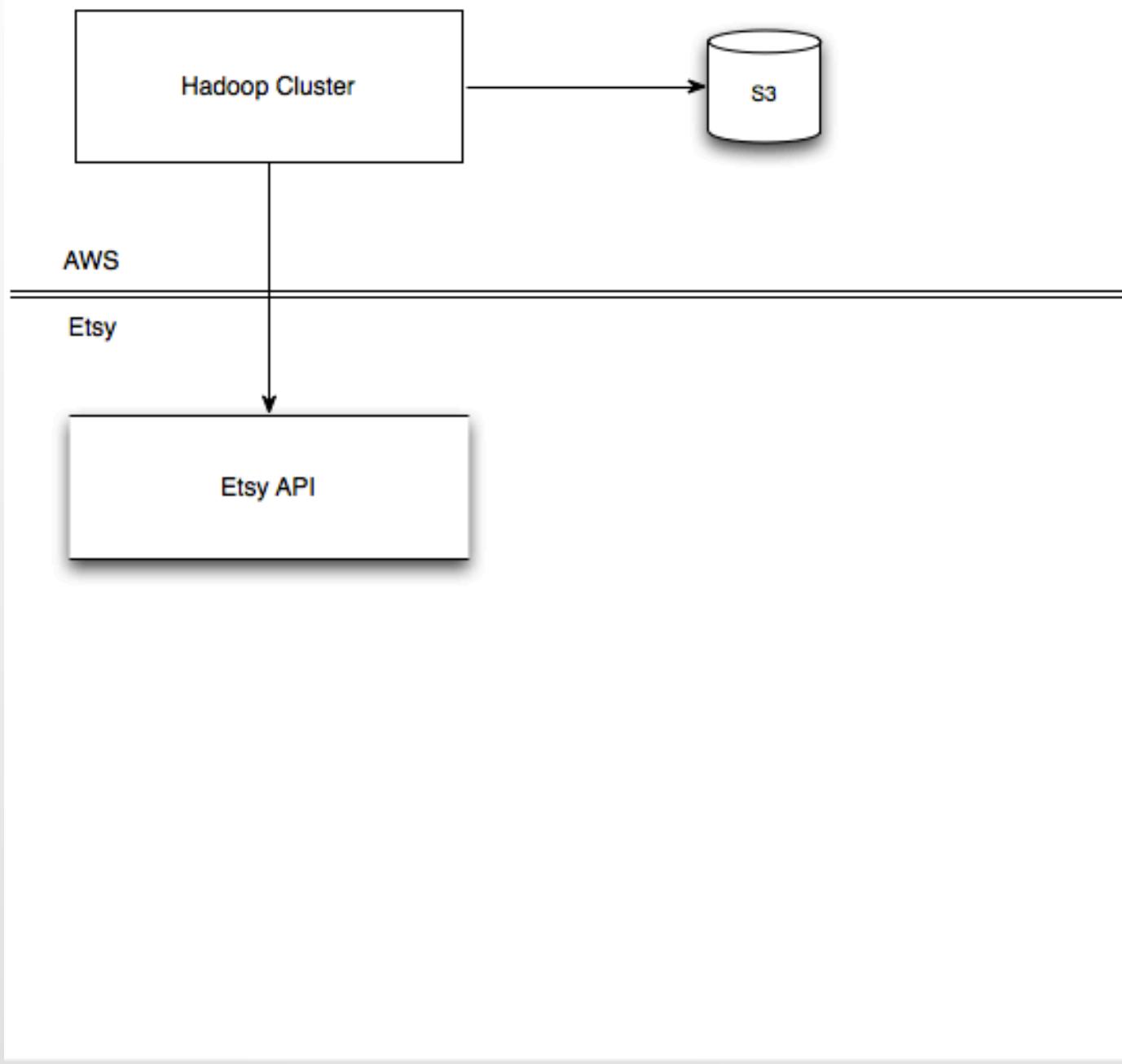
[6]

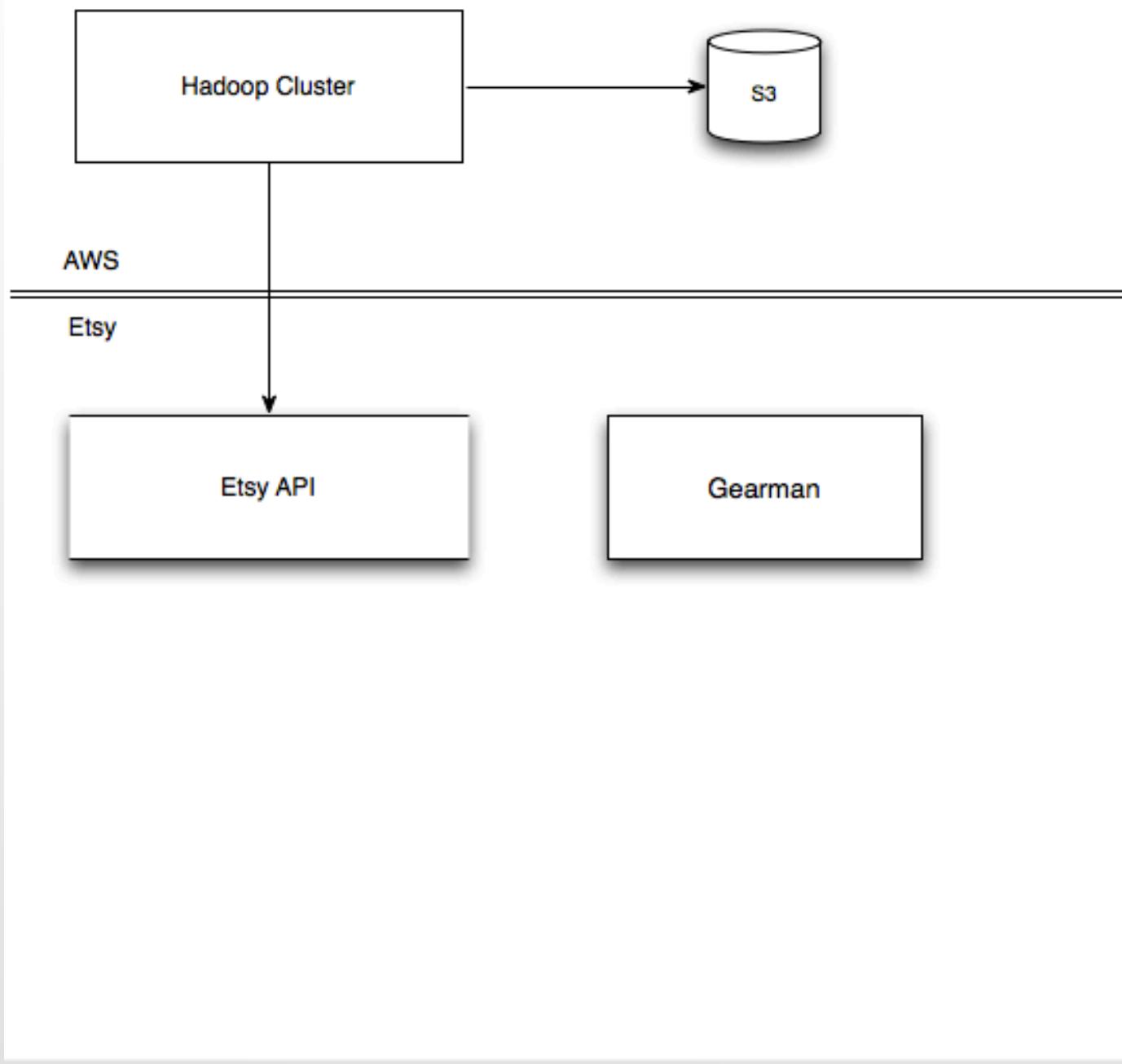
ad-hoc shell scripts
TSV into unsharded MySQL
not re-usable

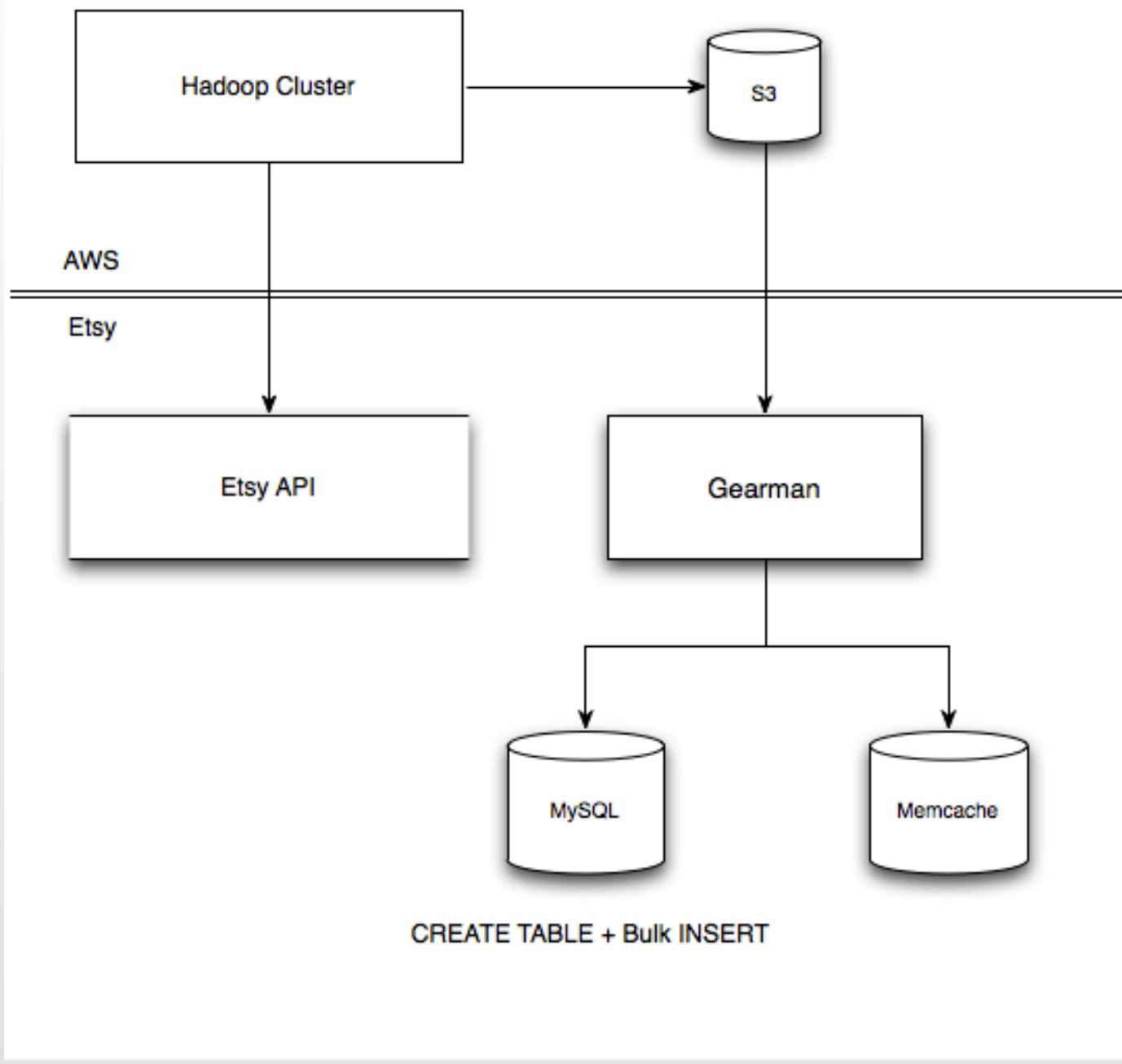
v2

Hadoop Cluster









datasets are versioned based upon
job execution time

Enqueue load

Target:

Bucket:

Path:

Time Path:

Pre-warm:

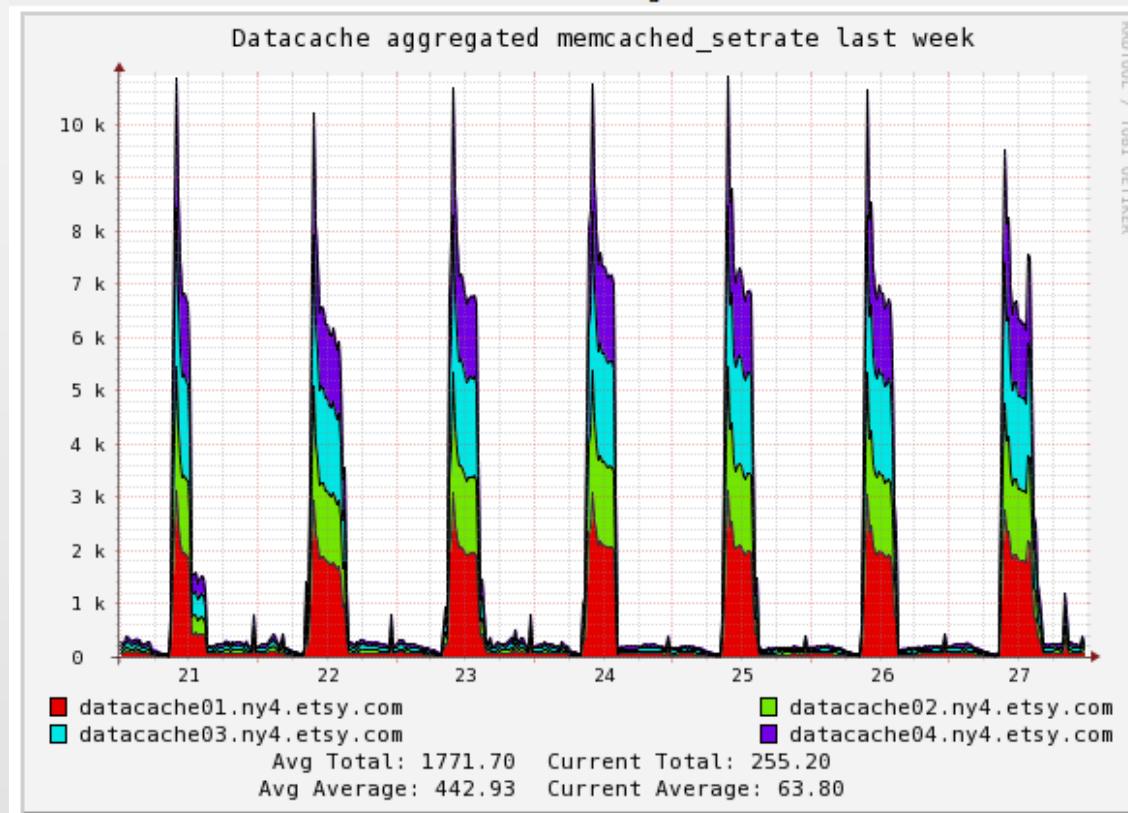
Loads

ID	Target	Time Path	Status	Total Jobs	Completed Jobs
265643610	Metrics_KeywordDistribution	2011_06_27/19_25_41	available	2	2
265622642	Metrics_StemmedTags	2011_06_27/19_27_28	available	2	2
265352260	Metrics_KeywordDistribution	2011_06_27/15_01_30	available	2	2
265332236	Metrics_StemmedTags	2011_06_27/15_03_24	available	2	2
266271519	Metrics_Shop	2011_06_27/13_00_02	available	2	2
265175322	UserSimilarity	2011_06_27/07_18_03	available	2	2
265109940	TasteTest_ItemTermIdf	2011_06_27/07_16_03	available	2	2
266169519	TasteTest_ShopTastemakerMale	2011_06_27/07_16_03	available	2	2
265103860	TasteTest_PathEdge	2011_06_27/07_16_03	available	2	2

MySQL Tables:

```
| tt_recommendable_favs____2011_06_24_07_16_03____262184924 |  
| tt_recommendable_favs____2011_06_25_07_16_03____263252724 |  
| tt_recommendable_favs____2011_06_26_07_16_03____264140496 |  
| tt_recommendable_favs____2011_06_27_07_16_03____266169499 |  
| tt_recommendation_batch |
```

Memcache Cluster:



Output dataset <-> ORM Model

PHP:

```
class EtsyModel_TasteTest_RecommendableFavoriteListing
    extends EtsyModel_DatasetRowBase {
```

PHP:

```
class EtsyModel_TasteTest_RecommendableFavoriteListing  
    extends EtsyModel_DatasetRowBase {
```

Cascading:

```
dataset "TasteTest_RecommendableFavoriteListing" do  
    project "user_id", "shop_id", "listing_id"  
    unique "user_id", "shop_id", "listing_id"  
    group_output_by "user_id"  
end
```

PHP:

```
class EtsyModel_TasteTest_RecommendableFavoriteListing  
    extends EtsyModel_DatasetRowBase {
```

Cascading:

```
dataset "TasteTest_RecommendableFavoriteListing" do  
  project "user_id", "shop_id", "listing_id"  
  unique "user_id", "shop_id", "listing_id"  
  group_output_by "user_id"  
end
```

PHP:

```
EtsyModel_Dataset::tap(  
  "TasteTest_RecommendableFavoriteListing"  
) ;
```

Old tables regularly dropped

how we're using this stack

analytics
(internal)

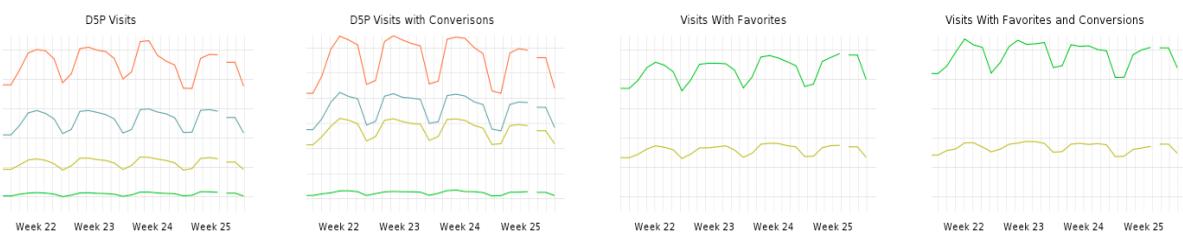
products
(external)

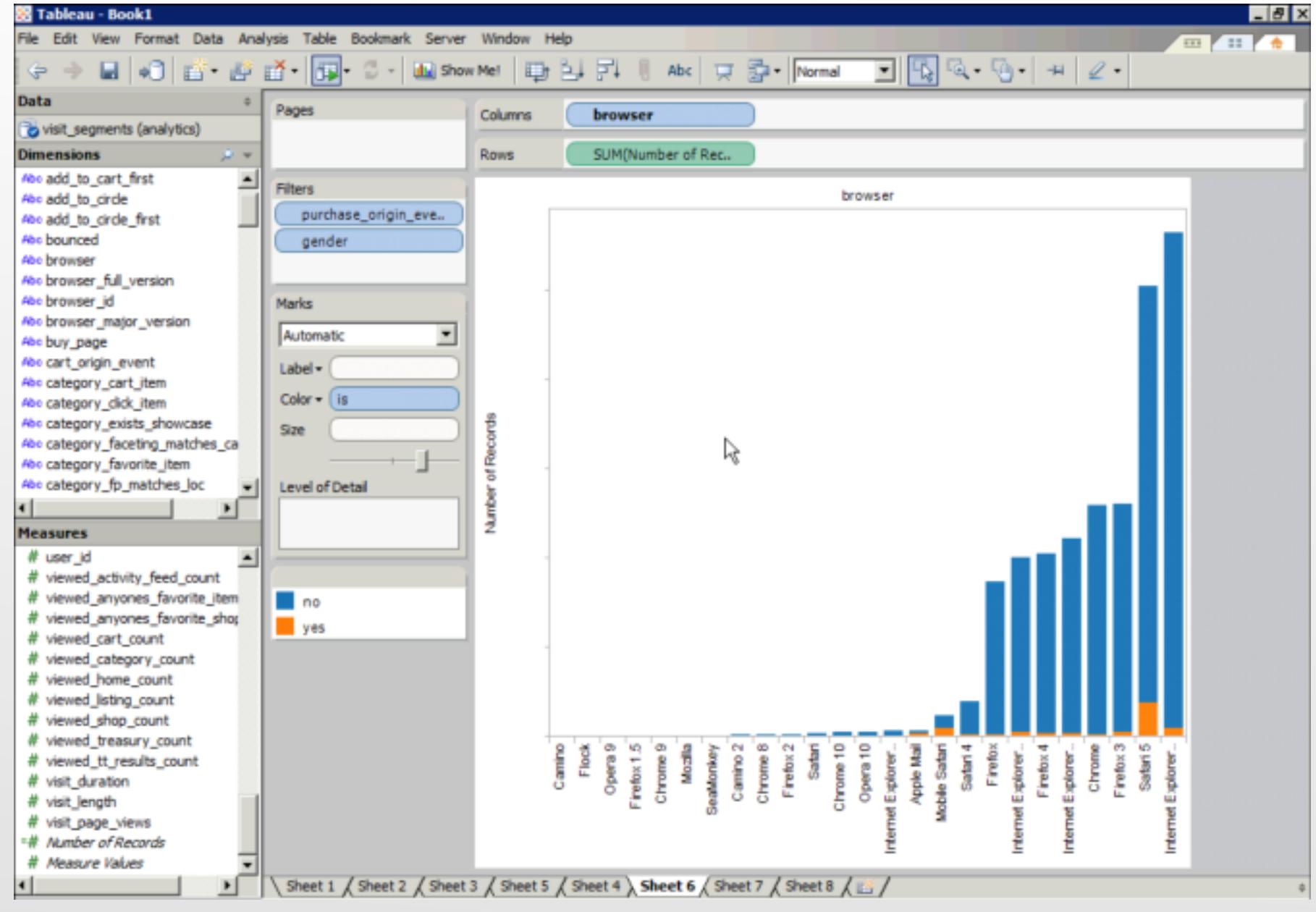
analytics

traffic (1 month)



visits (1 month)





products

Click the item that most suits your taste. [Show different items.](#)

Take the Taste Test!

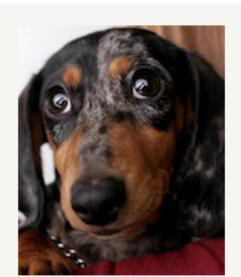
It's a fun way to explore more than 8.9 million items from shops around the world. Choose the item you like best, Etsy does the rest.

Buy
Sell
Community
Blog

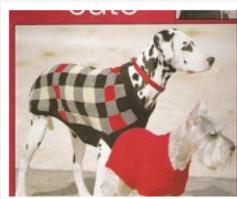
Etsy
205

Handmade

Gift Ideas for Dottie the Dog



Dottie likes **dog sweaters** [See more](#)



LEISURE ARTS Dog Sweaters
dislitr...



Friendly dog or pet sweat...
fiflime



Dog dish placemat S
JacksPup...

Greg Fodor's Favorites



gfodor
Greg Fodor
Brooklyn, United States

[Profile](#)

[Favorites](#)

Circles: 61

Feedback: 3, 100% pos.

[Contact](#)

[Shop](#)

[Items](#) [Shops](#) [Treasury Lists](#) [Suggested Shops](#)



MATTY8080



BeatUpCreations



mybeardedpigeon



search quality
recommendations

May 2011:
4,926 successful job runs



[5]

scale up from zero

isolation

isolation across runs

fresh machine each time

isolation between developers

no toe-stepping

heterogeneous clusters

big RAM when you need it
(but not when you don't)

need one machine?
use one machine.

writing jobs

The background image shows a dark, cloudy night sky with several bright, glowing lightning bolts striking down from the clouds. The horizon is dark and silhouetted against the bright sky.

PHENOMENAL COSMIC POWERS

prototyping

run slow, unoptimized version on 500 machine for < \$100

parameter tuning

Try N=1, 2, 5, 10 and see which results in best output



[9]



questions?



photo credits

- [1] by elfike <http://www.flickr.com/photos/elfike/157439707/>
- [2] by Dan4th <http://www.flickr.com/photos/43264265@N00/5371557240/>
- [3] by mandolux <http://www.flickr.com/photos/73935252@N00/34418046/>
- [4] by The Suss-Man <http://www.flickr.com/photos/8692813@N06/4580254188/>
- [5] by Stephen Rees <http://www.flickr.com/photos/60142746@N00/214461223/>
- [6] by Let Ideas Compete
http://www.flickr.com/photos/question_everything/3414827746/
- [7] by funkandjazz <http://www.flickr.com/photos/phunk/2484159004/>
- [8] by ViaMoi <http://www.flickr.com/photos/12187843@N07/3343619603/>
- [9] by kreg.steppe <http://www.flickr.com/photos/spyndle/500305000/>
- [10] clipart (really)
- [11] by Chris Pirillo <http://www.flickr.com/photos/49503157467@N01/34588230/>